Journal of Statistical Computation and Simulation
Vol. 76, No. 4, April 2006, 305–316

Taylor & Francis
Taylor & Francis Group

Estimating seasonal long-memory processes: a Monte Carlo study

VALDERIO A. REISEN*†, ALEXANDRE L. RODRIGUES† and WILFREDO PALMA‡

†Department of Statistics, CCE and PPGEA-CT-UFES, Vitoria – ES, Brazil
‡Department of Statistics, Pontificia Universidad Católica de Chile

(Revised 17 November 2003; in final form 20 October 2004)

This paper discusses extensions of the popular methods proposed by Geweke and Porter-Hudak [Geweke, J. and Porter-Hudak, S., 1983, The estimation and application of long memory times series models. *Journal of Time Series Analysis*, **4**(4), 221–238.] and Fox and Taqqu [Fox, R. and Taqqu, M.S., 1986, Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statistics*, **14**, 517–532.] for estimating the long-memory parameter of autoregressive fractionally integrated moving average models to the estimation of long-range dependent models with seasonal components. The proposed estimates are obtained from a selection of harmonic frequencies chosen between the seasonal frequencies. The maximum likelihood method given in Beran [Beran, J., 1994, *Statistic for Long-Memory Processes* (New York: Chapman & Hall).] and the semi-parametric approaches introduced by Arteche and Robinson [Arteche, J. and Robinson, P.M., 2000, Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis*, **21**(1), 1–25.] are also considered in the study. Our finite sample Monte Carlo investigations indicate that the proposed methods perform well and can be used as alternative estimating procedures when the data display both long-memory and cyclical behavior.

Keywords: Fractional differencing; Long-memory; Periodogram regression; Seasonality; Whittle maximum likelihood procedure

AMS Subject Classifications: 62M10, 62M15, 60G18

1. Introduction

Since their introduction by Granger and Joyeux [1] and Hosking [2], autoregressive fractionally integrated moving average (ARFIMA) processes have become very popular for modeling time series with long-memory behavior. A time series with this property has a slow and hyperbolically declining autocorrelation function or, equivalently, an infinite spectrum at zero frequency. A number of applied works have been published to illustrate the usefulness of ARFIMA models in different areas such as economy, hydrology, physics [see, *e.g.*, ref. [3] and references therein]. Moreover, recent books on time series analysis and econometrics have considered the ARFIMA model as one of their subjects, cf. Mills [4], Tsay [5] and Chan [6], among others.

*Corresponding author. Email: valderio@verao.cce.ufes.br

Journal of Statistical Computation and Simulation
ISSN 0094-9655 print/ISSN 1563-5163 online
http://www.tandf.co.uk/journals
DOI: 10.1080/10629360500107790

Since the 1980s, a number of estimators have been proposed for estimating the parameters of stationary and non-stationary ARFIMA processes. Most of them belong to either the semi-parametric or the parametric class. One of the most popular estimations methods in the first class is due to Geweke and Porter-Hudak [7] and it is based on the regression of the logarithm of the periodogram. In the second class, we may refer to the maximum likelihood methods proposed by Sowell [8] and Fox and Taqqu [9]. There are many works comparing and applying different estimators in the ARFIMA context including Hurvich and Ray [10], Velasco [11], Robinson [12], Reisen [13], Reisen *et al.* [14, 15] and Lopes *et al.* [16], among others.

The methodology for modeling time series with long-memory behavior has been recently extended to long-memory time series with seasonal components. Recent contributions related to the seasonal ARFIMA model (hereafter denoted by ARFISMA model) are Porter-Hudak [17], Hassler [18], Gray *et al.* [19, 20], Giraitis and Leipus [21], Woodward *et al.* [22], Arteche and Robinson [23] and Reisen *et al.* [24], among others.

The second order structure of seasonal long-memory time series is similar to the ARFIMA process, in the sense that the dependence between observations decay very slowly. However, the spectral density of an ARFIMA model has only one pole at zero frequency, whereas the spectral density of an ARFISMA process has singularities at the origin and at each seasonal frequency.

The main purpose of this work is to assess and compare the finite sample performance of several estimation methods in the context of long-memory time series exhibiting periodic behavior. As introduced in the following section, the process considered in this study is a particular case of the ARUMA model with two fractional memory parameters, defined in Giraitis and Leipus [21]. This result is shown in Lemma 1 of section 2. This paper continues a previous work by the authors, see ref. [24], dealing with the estimation of ARFISMA processes using modified forms of the Geweke and Porter-Hudak [7] method and the maximum likelihood estimator. This paper includes, in addition, the application of the parametric method proposed by Fox and Taqqu [9] to estimate ARFISMA models. The estimates of both the proposed methods are obtained from the evaluation of the periodogram at a number of harmonic frequencies chosen between the seasonal frequencies. The proposed estimation methodologies are compared with the semiparametric approaches proposed by Arteche and Robinson [23] and the maximum likelihood method [see, *e.g.*, section 5.3 of ref. [25]].

This paper is organized as follows. Section 2 discusses the seasonal long-memory model analyzed and the estimation procedures under study. Section 3 presents the results from a comparative Monte Carlo simulation analysis of the finite sample performance of these estimates. Final remarks are given in section 4.

2. Model and estimation methods

Let $\{X_t\}$ be a zero-mean ARFISMA$(p, d, q)\times(P, D, Q)_s$ process defined by

$$\varphi(B^s)\phi(B)(1-B)^d(1-B^s)^D X_t = \theta(B)\Theta(B^s)\varepsilon_t, \tag{1}$$

where d and D are real values, B is the lag operator, s is the seasonal period, $\Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$, $\Theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$, $\varphi(B^s) = 1 - \varphi_1 B^s - \varphi_{2s} B^{2s} - \cdots - \varphi_{Ps} B^{Ps}$ and $\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_{2s} B^{2s} - \cdots - \Theta_{Qs} B^{Qs}$ are polynomials of orders p, q, P, Q, respectively, with roots outside of the unit circle and $\{\varepsilon_t\}$ is a Gaussian white noise process with zero-mean and variance σ_ε^2. The fractional s difference is a generalization of the binomial expression $(1-B)^d$ and can be written as,

$$(1-B^s)^D = 1 - DB^s - \frac{D(1-D)B^{2s}}{2!} - \frac{D(1-D)(2-D)B^{3s}}{3!} - \cdots$$

For simplicity, in this paper, we restrict our attention to the case where $p = q = P = Q = 0$. In this situation, $\{X_t\}$ is an ARFISMA$(0, d, 0)\times(0, D, 0)_s$. When $d = 0$, $\{X_t\}$ is a rigid model in the terminology of Hassler [18].

Observe that Giraitis and Leipus [21] introduced the fractional autoregressive unit circle moving average ARUMA$(0, d_1, \ldots, d_r, 0)$ model with memory parameters $d_1, \ldots, d_r$ ($d_j \neq 0$, $j = 1, \ldots, r$) and fixed frequencies $0 \le \lambda_1 < \cdots < \lambda_r \le \pi$ as a stationary process Y_t obtained from the solution of the equation

$$\nabla^{d_1,\ldots,d_r}_{\lambda_1,\ldots,\lambda_r} Y_t = \varepsilon_t, \tag{2}$$

where ε_t is defined as previously, the parameters d_j are fractional degrees and $\nabla^{d_1,\ldots,d_r}_{\lambda_1,\ldots,\lambda_r} = \prod_{j=1}^{r}[(1-Be^{i\lambda_j})(1-Be^{-i\lambda_j})]^{d_j}$.

Assume for simplicity that the period s is even. Let λ_v be the seasonal frequency defined as $\lambda_v = 2\pi v/s$, for $v = 1, 2, \ldots, s/2$. The following result shows that the ARFISMA(0, d, 0) $\times$ $(0, D, 0)_s$ model can be written as an ARUMA process specified by equation (2) and also gives conditions on the parameters to assure causality and invertibility.

LEMMA 1 *Let X_t be defined as an ARFISMA$(0, d, 0) \times (0, D, 0)_s$ model. Then, X_t is an ARUMA$(0, d_0, d_j, \ldots, d_{s/2}, 0)$ process with $d_0 = (d + D)/2$ $(\lambda_0 = 0)$, $d_j = D$ for $j = 1, 2, \ldots, (s/2) - 1$ and $d_{s/2} = (D/2)(\lambda_{s/2} = \pi)$ and it is causal and invertible if and only if $|d + D| < 1/2$, $|D| < 1/2$.*

Proof The filter in model (1) may be written as,

$$\begin{aligned}(1-B)^d(1-B^s)^D &= (1-B)^d(1-B)^D(1+B)^D \prod_{j=1}^{(s/2)-1}[(1-Be^{i\lambda_j})(1-Be^{-i\lambda_j})]^D \\ &= [(1-Be^{i0})(1-Be^{-i0})]^{(d+D)/2}[(1-Be^{i\pi})(1-Be^{-i\pi})]^{D/2} \\ &\quad \times \prod_{j=1}^{(s/2)-1}[(1-Be^{i\lambda_j})]^D \\ &= \prod_{j=0}^{s/2}[(1-Be^{i\lambda_j})(1-Be^{-i\lambda_j})]^{d_j}. \end{aligned} \tag{3}$$

■

Thus, we conclude from Theorem 2 of Giraitis and Leipus [23] that equation (3) is the filter of an ARUMA(0, $(d+2)/2$, $D, \ldots, D/2$, 0) process. The conditions for causality and invertibility follow immediately from Theorem 1 of Giraitis and Leipus [21].

Note that for s odd the proof is similar, omitting the term $(1 + B)$ in equation (3). In this situation, the ARFISMA(0, d, 0) $\times$ $(0, D, 0)_s$ model is an ARUMA(0, $(d + D)/2$, $D, \ldots$, 0) model.

The spectral density of an ARFISMA(0, d, 0) × (0, D, 0)$_s$ process is given by

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi}\left[4^{-D}\sin^{-2D}\left(\frac{\lambda s}{2}\right)\right]\left[4^{-d}\sin^{-2d}\left(\frac{\lambda}{2}\right)\right] \tag{4}$$

for $0 \le \lambda \le \pi$. Thus, by allowing d and D to take positive and negative values, the spectral density of the ARFISMA(0, d, 0)×(0, D, 0)$_s$ may display zeroes or poles at some frequencies in the interval $(0, \pi)$.

Observe that at the seasonal frequencies $\lambda_v = 2\pi v/s$, $v = 0, 1, \ldots, [s/2]$, and for $D > 0$ and/or $d > 0$, the spectral density of the process $f(\lambda)$ becomes unbounded. Figures 1 and 2 illustrate the behavior of the spectral density for some values of D and d.

2.1 *Estimation methods*

In this section, we discuss briefly the application of several methods to the estimation of long-memory seasonal time series.

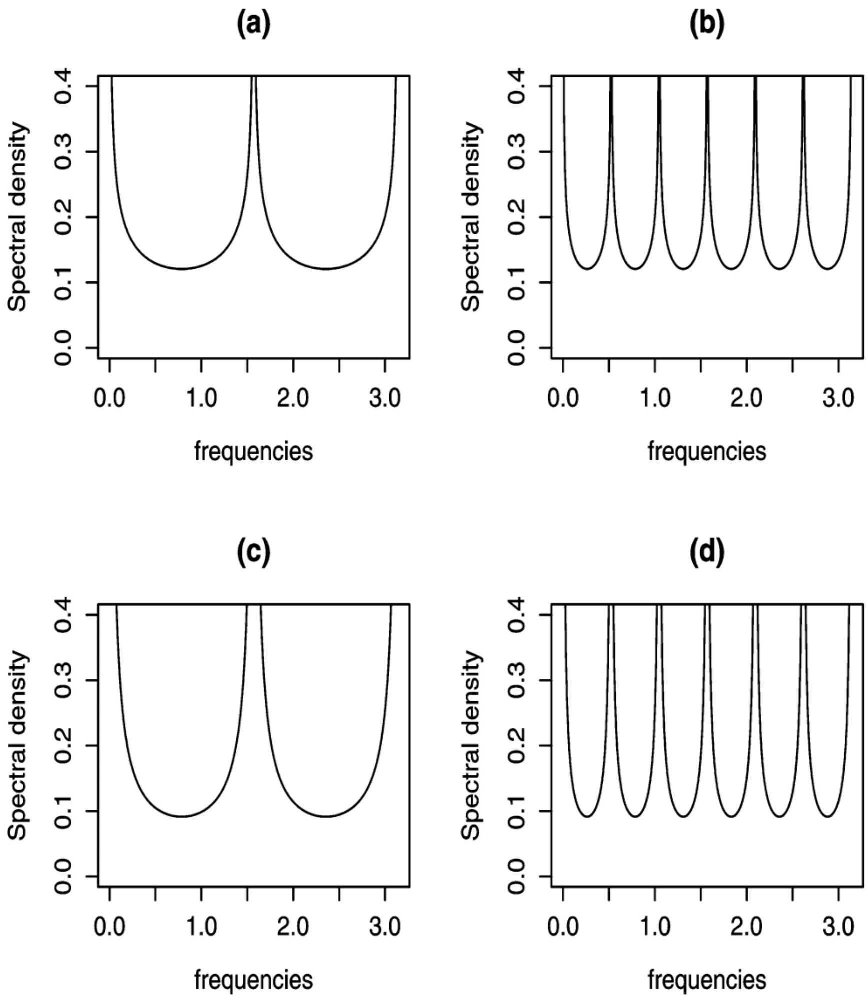

Figure 1. Spectral density of (a) Model ARFISMA(0, 0.2, 0)$_4$, (b) Model ARFISMA(0, 0.2, 0)$_{12}$, (c) Model ARFISMA(0, 0.4, 0)$_4$ and (d) Model ARFISMA(0, 0.4, 0)$_{12}$.

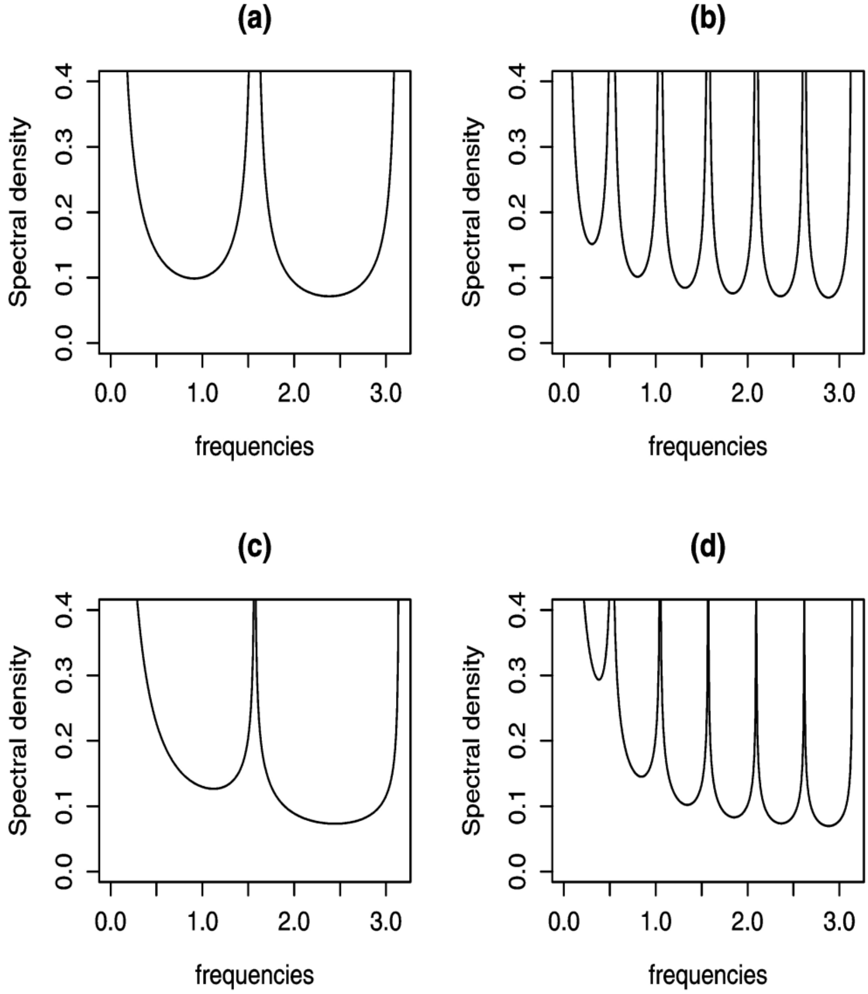

Figure 2. Spectral density of (a) Model ARFISMA(0, 0.2, 0) × (0, 0.4, 0)$_4$, (b) Model ARFISMA(0, 0.2, 0) × (0, 0.4, 0)$_{12}$, (c) Model ARFISMA(0, 0.4, 0) × (0, 0.2, 0)$_4$ and (d) Model ARFISMA(0, 0.4, 0) × (0, 0.2, 0)$_{12}$.

2.1.1 Regression methods. Let $I(\lambda) = n^{-1}|\sum_{t=1}^{n} X_t e^{i\lambda t}|^2$ be the periodogram function of the process. A multilinear regression equation is obtained by taking logarithms in expression (4),

$$\log f(\lambda) = \log \frac{\sigma_\varepsilon^2}{2\pi} - D \log \left[2 \sin\left(\frac{\lambda s}{2}\right)\right]^2 - d \log \left[2 \sin\left(\frac{\lambda}{2}\right)\right], \qquad (5)$$

for $0 \leq \lambda \leq \pi$. Estimates of d and D may be obtained by replacing $f(\lambda)$ by $I(\lambda)$ and then approximating the regression (5) by

$$\log I(\lambda) \cong a_0 - D \log \left[2 \sin\left(\frac{\lambda s}{2}\right)\right]^2 - d \log \left[2 \sin\left(\frac{\lambda}{2}\right)\right]^2 + U_\lambda, \qquad (6)$$

where a_0 is a constant and $U_\lambda = \ln I(\lambda)/f(\lambda) - E[\ln I(\lambda)/f(\lambda)]$.

Assume for simplicity that the number of observations n of $\{X_t\}$ is divisible by s. We consider the frequencies

$$\lambda_{v,j} = \frac{2\pi v}{s} + \frac{2\pi j}{n}, \quad v = 0, 1, \ldots, \left[\frac{s}{2}\right] - 1, \quad j = 1, 2, \ldots, m,$$

for some choice of the bandwidth m which satisfies the condition $(1/m) + (m/n) \to 0$ as n goes to ∞. Since λ must lie in the interval $(0, \pi)$, we set $\lambda_{[s/2],j} = \pi - (2\pi j/n)$. Different estimation methods for D and d may be obtained by appropriate choices of the harmonic frequency λ. Basically, the two regression estimators considered in this study are distinguished by the choice of the bandwidth m when regressing $\log(I(\lambda_{v,j}))$ on $\log[2\sin(\lambda_{v,j}s/2)]^2$ and $\log[2\sin(\lambda_{v,j}/2)]^2$. The regression methods proposed here are

(1) The estimator $\hat{d}_T$, T for total, produces the estimates by using all harmonic frequencies in the regression equation, that is, the regression is built from

$$\lambda_{v,j} = \frac{2\pi v}{s} + \frac{2\pi j}{n}, \quad v = 0, 1, \ldots, \left[\frac{s}{2}\right] - 1, \quad j = 1, 2, \ldots, m \text{ and } m = \left[\frac{n}{s}\right] - 1.$$

(2) Another estimator is the $\hat{d}_P$, P for partial, which is a particular case of $\hat{d}_T$ method. $\hat{d}_P$ considers a collection of the harmonic frequencies chosen around the right-hand side of the seasonal frequencies (left-hand side when $v = [s/2]$) and $m = [n/2s] - 1$. Observe that $m = n/2s$, $\lambda_{v,m}$ corresponds to half the distance between the seasonal frequencies λ_v and λ_{v+1}. It should be noted that in the rigid model ($d = 0$), the spectral density around each seasonal frequency is symmetric. Hence, the regression estimator $\hat{d}_P$ is similar to the regression $\hat{d}_T$ method but with fewer observations to be regressed. Consequently, it could be expected that $\hat{d}_T$ produce slightly more precise estimates than $\hat{d}_P$ method. On the other hand, $\hat{d}_P$ is computationally faster than $\hat{d}_T$.

Reisen *et al.* [24] also considered the regression equation restricted to a neighborhood of each seasonal frequency. Thus, regression estimates of the parameters can be obtained for each value of λ_v. However, this approach is omitted here because, as reported in that study, these estimators performed very poorly in finite sample simulations.

2.1.2 Fox–Taqqu method. This estimator, hereafter denoted by $\hat{d}_{\mathrm{W}}$, is a parametric procedure due to Fox and Taqqu [9] and Whittle [26] for Gaussian long-memory processes and it is based on the periodogram and the spectral density functions. The subscript W stands for Whittle, who proposed the following likelihood function for parameter estimation in the context of short-memory time series. This estimator is obtained by using all harmonic frequencies between the seasonal frequencies as in ($\hat{d}_T$). It is calculated by minimizing the approximate Gaussian log-likelihood

$$\mathcal{L}_{\mathrm{W}}(\boldsymbol{\theta}) = \frac{1}{2n} \sum_j \left\{ \ln f_{\boldsymbol{\theta}}(\lambda_j) + \frac{I(\lambda_j)}{f_{\boldsymbol{\theta}}(\lambda_j)} \right\}, \tag{7}$$

where $f_{\boldsymbol{\theta}}$ is the spectral density,

$$\boldsymbol{\theta} = (d, D, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, \varphi_1, \ldots, \varphi_P, \Theta_1, \ldots, \Theta_Q, \sigma_\varepsilon^2),$$

denotes the vector of unknown parameters and $\sum_j$ is sum over $j = 1, \ldots, n-1$, excluding those values λ_j coinciding with the seasonal frequencies. In this work, the spectral density $f_{\boldsymbol{\theta}}$ is specified by equation (4).

2.1.3 The semi-parametric $\hat{d}_v$ and $\hat{d}_Q$ methods. The semi-parametric estimation methods proposed in Arteche and Robinson [23] are denoted in this work by $\hat{d}_v$ for $v = 0, \ldots, s/2$ and $\hat{d}_Q$. These estimators correspond to methods described in section 3 and 4 of Arteche and Robinson [23], respectively. As these estimates have been well described in the previous reference, we omit the formulas here.

The bandwidth, m, used by these estimators is the same as the one chosen for the estimate $\hat{d}_P$, *i.e.*, $m = [n/2s] - 1$. Note that for the frequencies in $(0, \pi)$, the $\hat{d}_v$ estimators use the information in both sides of the seasonal frequency λ_v. Thus, $\hat{d}_v$ is the mean value of the left and right log-regression estimators. For more details about this technique, see equation (3.4) of ref. [23].

2.1.4 Maximum likelihood method. The maximum likelihood estimator ($\hat{d}_{\mathrm{ML}}$) is obtained by maximizing

$$\mathcal{L}_n(\boldsymbol{\theta}) = -\frac{1}{2} \log \det T(f_{\boldsymbol{\theta}}) - \frac{1}{2} \mathbf{X}' T(f_{\boldsymbol{\theta}})^{-1} \mathbf{X}, \tag{8}$$

where $\boldsymbol{\theta} = (d, D)$ is the parameter vector, $f_{\boldsymbol{\theta}}$ is the spectral density, $\mathbf{X} = (x_1, \ldots, X_n)'$ and $T(f_{\boldsymbol{\theta}})$ is the variance–covariance matrix of the Gaussian process $\{X_t\}$. A detailed revision of this approach is found in section 5.3 of ref. [25].

3. Monte Carlo simulation study

In this section, we study the finite sample performance of the methods discussed in section 2 via Monte Carlo experiments. We carried out several simulations for different combinations of parameters, seasonal period and sample sizes. All the results shown in this section are based on 1000 replications.

The sample mean and the mean squared error (MSE) are presented in tables 1–6. The calculations were carried out by means of an Ox program in an AMD Athlon XP 1800 computer. We considered sample sizes $n = 240$, $n = 480$ and $n = 3600$, and seasonal periods $s = 4$ and $s = 12$. The models and the values of the long-memory parameters D and d are specified in each table. Simulations with other values of d and D gave similar results and they are available upon request. The ARFISMA processes were simulated following the method suggested by Hosking [27] with Gaussian noise with unit variance. Tables 1–3 present the results concerned to the rigid model ARFISMA$(0, D, 0)_s$ and the remaining tables display the results related to the ARFISMA$(0, d, 0) \times (0, D, 0)_s$ process.

3.1 *Estimation of the ARFISMA(0, D, 0)$_s$ model*

It can be seen from tables 1–3 that the methods $\hat{d}_T$ and $\hat{d}_P$ have good performance (small bias and MSE), even for small sample size. It seems that better estimates are obtained from $\hat{d}_T$ than $\hat{d}_P$, as the former method involves all harmonic frequencies in the regression equation. The Whittle estimator $\hat{d}_{\mathrm{W}}$ produces larger bias than $\hat{d}_T$ and $\hat{d}_P$. However, as n increases the bias of $\hat{d}_{\mathrm{W}}$ decreases substantially, and these three methods become very competitive. The performance of these methods is directly related to the fact that the spectral density is symmetric around all seasonal frequencies (figure 1).

Comparing the two parametric methods, $\hat{d}_{\mathrm{W}}$ and $\hat{d}_{\mathrm{ML}}$, the former seems to be more biased. However, their results are fairly similar for large n.

Table 1. Estimates of D when $D = 0.2$ and $s = 4$.

n	Statistics	Estimators $\hat{d}_T$	$\hat{d}_P$	$\hat{d}_0$	$\hat{d}_1$
240	Mean	0.2003	0.2034	0.2225	0.1962
	MSE	0.0059	0.0079	0.0321	0.0110
480	Mean	0.2026	0.2002	0.2011	0.2013
	MSE	0.0030	0.0034	0.0136	0.0050
3600	Mean	0.1970	0.1981	0.1989	0.1952
	MSE	0.0002	0.0003	0.0012	0.0005
n	Statistics	$\hat{d}_2$	$\hat{d}_Q$	$\hat{d}_{\mathrm{W}}$	$\hat{d}_{\mathrm{ML}}$
240	Mean	0.1845	0.1505	0.1644	0.1847
	MSE	0.0230	0.0212	0.0033	0.0039
480	Mean	0.2038	0.1691	0.1780	0.1964
	MSE	0.0119	0.0092	0.0016	0.0016
3600	Mean	0.1988	0.1930	0.1944	0.2067
	MSE	0.0010	0.0009	0.0002	0.0003

Table 2. Estimates of D when $D = 0.2$ and $s = 4$.

n	Statistics	Estimators $\hat{d}_T$	$\hat{d}_P$	$\hat{d}_0$	$\hat{d}_1$
240	Mean	0.4047	0.4087	0.4301	0.3983
	MSE	0.0054	0.0074	0.0300	0.0112
480	Mean	0.4063	0.4039	0.4044	0.4045
	MSE	0.0031	0.0034	0.0153	0.0052
3600	Mean	0.3988	0.3999	0.4006	0.3971
	MSE	0.0002	0.0003	0.0013	0.0005
n	Statistics	$\hat{d}_2$	$\hat{d}_Q$	$\hat{d}_{\mathrm{W}}$	$\hat{d}_{\mathrm{ML}}$
240	Mean	0.3902	0.3559	0.3470	0.3891
	MSE	0.0235	0.0204	0.0065	0.0040
480	Mean	0.4084	0.3751	0.3744	0.4038
	MSE	0.0123	0.0093	0.0022	0.0018
3600	Mean	0.4005	0.3948	0.3951	0.4281
	MSE	0.0011	0.0009	0.0002	0.0011

Table 3. Estimates of D when $D = 0.2$ and $s = 12$.

n	Statistics	Estimators $\hat{d}_T$	$\hat{d}_P$	$\hat{d}_0$	$\hat{d}_1$	$\hat{d}_2$	$\hat{d}_3$
240	Mean	0.1908	0.1966	0.1864	0.1737	0.2236	0.1875
	MSE	0.0093	0.0165	0.1532	0.0685	0.0672	0.0625
480	Mean	0.2005	0.1972	0.1597	0.2121	0.1900	0.1946
	MSE	0.0035	0.0062	0.0586	0.0211	0.0278	0.0214
3600	Mean	0.1979	0.1994	0.2000	0.1948	0.1951	0.1953
	MSE	0.0003	0.0005	0.0045	0.0020	0.0020	0.0020
n	Statistics	$\hat{d}_4$	$\hat{d}_5$	$\hat{d}_6$	$\hat{d}_Q$	$\hat{d}_{\mathrm{W}}$	$\hat{d}_{\mathrm{ML}}$
240	Mean	0.1749	0.2081	0.1393	−0.0391	0.1290	0.1848
	MSE	0.0748	0.0598	0.1818	0.1402	0.0058	0.0060
480	Mean	0.2012	0.2214	0.1898	0.0593	0.1507	0.1965
	MSE	0.0186	0.0195	0.0403	0.0587	0.0031	0.0020
3600	Mean	0.1998	0.2013	0.2021	0.1806	0.1903	0.2055
	MSE	0.0022	0.0020	0.0029	0.0032	0.0003	0.0002

On the other hand, the semi-parametric methods $\hat{d}_v$ ($v = 0, 1, \ldots, s/2$) and $\hat{d}_Q$ produce estimates that become competitive with the regression methods $\hat{d}_T$ and $\hat{d}_P$ for large sample size.

For all estimators, the bias is more noticeable when $s = 12$ but decreases as n increases (table 3). For small sample sizes, the MSE of all the regression methods increases with s beause of the fact that fewer number of harmonic frequencies are used. This effect becomes insignificant as n increases. Note that the asymptotic variance of the parameters for long-memory Gaussian processes is $\Gamma(\boldsymbol{\theta})^{-1}/n$ where $\Gamma(\boldsymbol{\theta})$ is the Fisher information matrix given by, cf. Dahlhaus [28]:

$$\Gamma(\boldsymbol{\theta}) = \frac{1}{4\pi}\int_{-\pi}^{\pi}[\nabla \log f(\lambda)][\nabla \log f(\lambda)]' \mathrm{d}\lambda. \tag{9}$$

For the ARFIMA(0, D, 0)$_s$ model, the asymptotic variance of $\hat{D}$ does not depend on the seasonal period s nor the value of D. This asymptotic variance is $\sim$ 0.0025, 0.0013 and 0.0002 for $n = 240$, $n = 480$ and $n = 3600$, respectively. Thus, the proposed estimators $\hat{d}_T$ and $\hat{d}_P$ have asymptotic variances close to the optimal values for $n = 3600$.

3.2 *Estimation of the ARFISMA(0, d, 0) × (0, D, 0)$_s$ model*

Tables 4–6 display the estimation results when d is also included in the model and the process is not restricted only to the stationary conditions. That is, we also consider the case

Table 4. Estimates of d and D when $d = 0.1$, $D = 0.3$ and $s = 4$.

		Estimators			
n	Statistics	$\hat{d}_T$	$\hat{d}_P$	$\hat{d}_0$	$\hat{d}_1$
240	Mean$_d$	0.1051	0.1042	0.1127	–
	MSE	0.0054	0.0054	2.1947	–
	Mean$_D$	0.3010	0.3040	0.3167	0.2953
	MSE	0.0063	0.0084	2.8071	0.0748
	corr($\hat{d}$, $\hat{D}$)	−0.2773	−0.2717	−0.9948	–
480	Mean$_d$	0.0960	0.0970	0.2358	–
	MSE	0.0029	0.0031	1.3140	–
	Mean$_D$	0.3057	0.3036	0.1544	0.2999
	MSE	0.0033	0.0041	1.6534	0.0198
	corr($\hat{d}$, $\hat{D}$)	−0.2911	−0.3849	−0.9953	–
3600	Mean$_d$	0.103	0.10256	0.1036	–
	MSE	0.0003	0.0003	0.1153	–
	Mean$_D$	0.2973	0.2982	0.2968	0.2992
	MSE	0.0003	0.0004	0.1404	0.0018
	corr($\hat{d}$, $\hat{D}$)	−0.3194	−0.3187	−0.9954	–

n	Statistics	$\hat{d}_2$	$\hat{d}_{\mathrm{W}}$	$\hat{d}_{\mathrm{ML}}$
240	Mean$_d$	–	0.0970	0.0792
	MSE	–	0.0034	0.0030
	Mean$_D$	0.2894	0.2495	0.2869
	MSE	0.0537	0.0059	0.0029
	corr($\hat{d}$, $\hat{D}$)	–	−0.2564	−0.2736
480	Mean$_d$	–	0.0990	0.0840
	MSE	–	0.0017	0.0018
	Mean$_D$	0.2908	0.2727	0.2931
	MSE	0.0245	0.0022	0.0016
	corr($\hat{d}$, $\hat{D}$)	–	−0.2696	−0.2803
3600	Mean$_d$	–	0.10158	0.0958
	MSE	–	0.0002	0.0002
	Mean$_D$	0.2956	0.2940	0.3023
	MSE	0.0017	0.0002	0.0002
	corr($\hat{d}$, $\hat{D}$)	–	−0.3003	−0.2906

$|d + D| > 1/2$. Apart from the mean and the MSE of the estimates, these tables also show the sample correlation between the estimates of d and D for the estimators $\hat{d}_T$, $\hat{d}_P$, $\hat{d}_W$ and $\hat{d}_{ML}$. Because of the large bias produced by $\hat{d}_Q$, this method was omitted in these tables. Besides, the use of $\hat{d}_{ML}$ is only restricted to the stationary case. It should be noted that the estimators $\hat{d}_v$, for $v = 1, \ldots, [s/2]$, are used for estimating D only as its seasonal frequency is away from zero. The estimator $\hat{d}_v$ combines the effect of d and D, as these parameters cannot be identified separately, see for example, the high value of sample correlation between the estimates produced by $\hat{d}_0$ in table 5. This estimator produces biased estimates of both memory parameters and displays a substantially large MSE. However, as v increases the bias and the MSE of the estimate of D decreases.

Observe that overall, estimators $\hat{d}_T$ and $\hat{d}_P$ display a good performance. Assuming that the process is Gaussian, Reisen *et al.* [24] obtained an asymptotic formula for correlation of the estimates of d and D. For $s = 4$, corr($\hat{d}$, $\hat{D}$)$= -0.247$ and for $s = 12$, corr($\hat{d}$, $\hat{D}$)$= -0.083$. The sample correlations of the proposed estimators are very close to the one given by the ML estimate. For small sample size, the parametric estimators, $\hat{d}_W$ and $\hat{d}_{ML}$, have larger bias than the semi-parametric methods, $\hat{d}_T$ and $\hat{d}_P$ (table 4).

For the non-stationary case (table 5), the results are fairly similar to the previous cases. The estimator $\hat{d}_T$ seems to have the best results, and for large n, the estimators $\hat{d}_T$, $\hat{d}_P$ and $\hat{d}_W$ yield very close results.

Table 5. Estimates of d and D when $d = 0.2$, $D = 0.4$ and $s = 4$.

		Estimators		
n	Statistics	$\hat{d}_T$	$\hat{d}_P$	$\hat{d}_0$
240	Mean$_d$	0.2152	0.2138	0.1564
	MSE	0.0053	0.0052	2.2226
	Mean$_D$	0.4043	0.4090	0.4991
	MSE	0.0061	0.0080	2.8466
	corr($\hat{d}$, $\hat{D}$)	−0.3446	−0.3028	−0.9954
480	Mean$_d$	0.2056	0.2066	0.3661
	MSE	0.0032	0.0035	1.3501
	Mean$_D$	0.4078	0.4054	0.2350
	MSE	0.0033	0.0041	1.6973
	corr($\hat{d}$, $\hat{D}$)	−0.2378	−0.3205	−0.9950
3600	Mean$_d$	0.20499	0.20452	0.1036
	MSE	0.0004	0.0004	0.1153
	Mean$_D$	0.3989	0.3999	0.4184
	MSE	0.0003	0.0004	0.1380
	corr($\hat{d}$, $\hat{D}$)	−0.2684	−0.2809	−0.9953
n	Statistics	$\hat{d}_1$	$\hat{d}_2$	$\hat{d}_W$
240	Mean$_d$	–	–	0.2064
	MSE	–	–	0.0038
	Mean$_D$	0.4025	0.3925	0.3442
	MSE	0.0666	0.0521	0.0073
	corr($\hat{d}$, $\hat{D}$)	–	–	−0.3113
480	Mean$_d$	–	–	0.2077
	MSE	–	–	0.0021
	Mean$_D$	0.4052	0.3960	0.3724
	MSE	0.0197	0.0257	0.0024
	corr($\hat{d}$, $\hat{D}$)	–	–	−0.2521
3600	Mean$_d$	c	–	0.20421
	MSE	–	–	0.0002
	Mean$_D$	0.4012	0.3966	0.3953
	MSE	0.0018	0.0018	0.0002
	corr($\hat{d}$, $\hat{D}$)	–	–	−0.2820

Table 6. Estimates of d and D when $d = 0.1$, $D = 0.3$ and $s = 12$.

		Estimators					
n	Statistics	$\hat{d}_T$	$\hat{d}_P$	$\hat{d}_0$	$\hat{d}_1$	$\hat{d}_2$	$\hat{d}_3$
240	Mean_d	0.1038	0.1107	0.0003	–	–	–
	MSE	0.0053	0.0075	7.5976	–	–	–
	Mean_D	0.2931	0.2974	0.4078	0.2604	0.2821	0.2771
	MSE	0.0092	0.0161	10.2740	0.4123	0.3588	0.4472
	$\text{corr}(\hat{d}, \hat{D})$	−0.0673	−0.0362	−0.9928	–	–	–
480	Mean_d	0.0958	0.0955	0.2611	–	–	–
	MSE	0.0026	0.0038	4.4582	–	–	–
	Mean_D	0.3038	0.3016	0.0823	0.3495	0.2426	0.3259
	MSE	0.0037	0.0066	5.9931	0.1037	0.1164	0.1270
	$\text{corr}(\hat{d}, \hat{D})$	−0.0623	−0.1443	−0.9944	–	–	–
3600	Mean_d	0.1023	0.1006	0.0728	–	–	–
	MSE	0.0003	0.0003	0.3517	–	–	–
	Mean_D	0.2991	0.3006	0.3334	0.2873	0.2907	0.3055
	MSE	0.0003	0.0005	0.4195	0.0056	0.0052	0.0062
	$\text{corr}(\hat{d}, \hat{D})$	−0.1061	−0.1128	−0.9946	–	–	–
n	Statistics	$\hat{d}_4$	$\hat{d}_5$	$\hat{d}_6$	$\hat{d}_W$	$\hat{d}_{\text{ML}}$	
240	Mean_d	–	–	–	0.0920	0.0792	
	MSE	–	–	–	0.0034	0.0030	
	Mean_D	0.3446	0.4316	0.20516	0.1808	0.2904	
	MSE	0.3884	0.4735	0.52177	0.0154	0.0032	
	$\text{corr}(\hat{d}, \hat{D})$	–	–	–	0.0634	−0.1525	
480	Mean_d	–	–	–	0.0963	0.0864	
	MSE	–	–	–	0.0016	0.0016	
	Mean_D	0.2901	0.3313	0.29617	0.2276	0.2951	
	MSE	0.1102	0.1096	0.12614	0.0064	0.0013	
	$\text{corr}(\hat{d}, \hat{D})$	–	–	–	0.0029	−0.0668	
3600	Mean_d	–	–	–	0.1010	0.0959	
	MSE	–	–	–	0.0002	0.0002	
	Mean_D	0.3177	0.3126	0.30844	0.2888	0.3027	
	MSE	0.0067	0.0063	0.00851	0.0003	0.0002	
	$\text{corr}(\hat{d}, \hat{D})$	–	–	–	−0.0788	−0.0906	

With the aim of investigating the effect of s on the estimates of both memory parameters and fixed sample size n, one stationary case was simulated and it is presented in table 6. From this investigation, the estimators $\hat{d}_T$ and $\hat{d}_P$ seem to produce very good estimates. For large sample size, these two periodogram-based estimators and the parametric estimators are very competitive.

4. Conclusions

Our empirical Monte Carlo investigation indicates that the two estimators proposed in this work, namely, $\hat{d}_T$ and $\hat{d}_P$ perform well even for small sample sizes. Consequently, they are reasonable methods for dealing with seasonal long-range dependent data. These two regression methods proposed are natural extensions of the one proposed originally by Geweke and Porter-Hudak [7] for non-seasonal long-memory processes. These periodogram-based methods have the advantage that they are not only restricted to the stationary case but they can also be used for seasonal non-stationary ARFIMA model. The regression methods $\hat{d}_T$ and $\hat{d}_P$ can be easily implemented. However, the Whittle estimator $\hat{d}_{\text{W}}$ requires the implementation of a high precision numerical procedure in order to minimize the quasi-likelihood function (7). Finally, in this work, we also compared the finite sample performance of the periodogram-based methods with the semi-parametric methodology proposed by Arteche and Robinson [23] and the maximum likelihood approach. Other estimation methods of the long-memory parameter such as that given in Reisen [13] are still under research by the first author in the context of seasonal unit root and long-memory processes.

Acknowledgements

The first two authors gratefully acknowledge partial financial support from PIBIC-UFES and CNPq/Brazil. The third author was partially supported by Fondecyt grant number 1040934. We also would like to thank Ricardo Olea for carrying out part of the Monte Carlo experiments, and the referee of the JSCS for his contribution to the improvement of the paper.

References

[1] Granger, W.J. and Joyeux, R., 1980, An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, **1**(1), 5–29.
[2] Hosking, J., 1981, Fractional differencing. *Biometrika*, **68**(1), 165–176.
[3] Doukhan, P., Oppenheim, G. and Taqqu, M.S., 2003, *Theory and Applications of Long-Range Dependence* (Boston: Birkhäuser Boston Inc.).
[4] Mills, T.C., 1997, *The Ecometric Modelling of Financial Time Series* (London: Cambridge University Press).
[5] Tsay, R.S., 2002, *Analysis of Financial Time Series* (New York: Wiley).
[6] Chan, N.H., 2002, *Time Series. Applications to Finance* (New York: Wiley).
[7] Geweke, J. and Porter-Hudak, S., 1983, The estimation and application of long memory times series models. *Journal of Time Series Analysis*, **4**(4), 221–238.
[8] Sowell, F., 1992, Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics*, **53**, 165–188.
[9] Fox, R. and Taqqu, M.S., 1986, Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statistics*, **14**, 517–532.
[10] Hurvich, C.M. and Ray, B.K., 1995, Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes. *Journal of Time Series Analysis*, **16**, 017–042.
[11] Velasco, C., 1999, Gaussian semiparametric estimation of non-stationary time series. *Journal of Time Series Analysis*, **20**, 87–127.
[12] Robinson, P.M., 1995, Log-Periodogram regression of time series with long range dependence. *Annals of Statistics*, **23**(3), 1048–1072.
[13] Reisen, V.A., 1994, Estimation of the fractional difference parameter in the ARFIMA(p,d,q) model using the smoothed periodogram. *Journal of Time Series Analysis*, **15**(1), 335–350.
[14] Reisen, V., Abraham, B. and Toscano, E.M.M., 2000, Parametric and semiparametric estimation of stationary univariate ARFIMA models. *Brazilian Journal of Probability and Statistics*, **14**, 185–206.
[15] Reisen, V., Abraham, B. and Lopes, S., 2001, Estimation of parameters in ARFIMA processes: a simulation study. *Communications in Statistics*, **30**(4), 787–803.
[16] Lopes, S.R.C., Olbermann, B.P. and Reisen, V.A., 2004, A comparison of estimation methods in non-stationary ARFIMA processes. *Journal of Statistical Computation and Simulation*, **74**(5), 339–347.
[17] Porter-Hudak, N., 1990, An aplication of the seasonal fractionally differenced model to the monetary aggegrates. *Journal of American Statistical Association*, **85**(410), 338–344.
[18] Hassler, U., 1994, (Mis)specification of long memory in seasonal time series. *Journal of Time Series Analysis*, **15**(1), 19–30.
[19] Gray, H.L., Zhang, N.F. and Woodward, W.A., 1989, On generalized fractional processes. *Journal of Time Series Analysis*, **10**(3), 233–257.
[20] Gray, H.L., Zhang, N.F. and Woodward, W.A., 1994, Correction to on generalized fractional processes. *Journal of Time Series Analysis*, **15**(5), 561–562.
[21] Giraitis, L. and Leipus, R., 1995, A generalized fractionally differencing approach in long memory modelling. *Lithuanian Mathematical Journal*, **35**, 53–65.
[22] Woodward, W.A., Cheng, Q.C. and Gray, H.L., 1998, A k-factor GARMA long-memory model. *Journal of Time Series Analysis*, **19**(4), 485–504.
[23] Arteche, J. and Robinson, P.M., 2000, Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis*, **21**(1), 1–25.
[24] Reisen, V., Rodrigues, A. and Palma, W., 2004, Estimation of seasonal fractionally integrated processes. *Computational Statistics and Data Analysis*, **50**, 568–582.
[25] Beran, J., 1994, *Statistic for Long-Memory Processes* (New York: Chapman and Hall).
[26] Whittle, P., 1953, Estimation and information in stationary time series. *Arkiv for Matematik*, **2**, 423–434.
[27] Hosking, J., 1984, Modelling persistence in hydrological time series using fractional differencing. *Water Resources Research*, **20**(12), 1898–1908.
[28] Dahlhaus, R., 1989, Efficient parameter estimation of self-similar processes. *Annals of Statistics*, **17**, 1749–1766.